

CLAIMS

1. A method of encoding a document image, the method comprising:
extracting one or more picture areas from the document image;
5 extracting one or more character areas from the document image;
obtaining a background image by subtracting the image and character areas
from the document image;
classifying character blocks of the character areas with reference to dynamically
generated templates; and
10 encoding the background image utilising a SAQ wavelet encoder.

2. The method as claimed in claim 1, wherein the extracting of the picture
areas and/or the character areas comprises marking blocks partitioned from the
document image based on features of wavelet coefficients of the respective blocks.

3. The method as claimed in claims 1 or 2, wherein the extracting of the
pictures areas comprises a hierarchical extraction comprising extracting picture blocks
from the document image to generate one or more initial picture areas and refining the
initial picture areas by extracting picture pixels adjacent to the initial picture areas.

4. The method as claimed in any one of claims 1 to 3, wherein the extracting
of the character areas from the document image comprises utilising a customised
definition of the connectivity of the pixels.

5. The method as claimed in any one of claims 1 to 4, further comprising
generating style data as a description of the templates and character blocks.

6. The method as claimed in claim 5, wherein the classifying the character
blocks comprises a hierarchical matching comprising matching the style of each
character block based on the style data and then matching each character block against
selected ones of the templates based on the style data matching.

7. The method as claimed in any one of claims 1 to 6, wherein the
classifying of the character blocks based on the templates comprises morphological
matching.

8. The method as claimed in claim 7, wherein the morphological matching comprises matching algorithms M_1 and M_2 ,

5 9. The method as claimed in claim 8, wherein different structure elements are utilised for different types of document images.

10 10. The method as claimed in any one of claims 1 to 9, further comprising bit plane storage of a compressed stream of the document image in the order of character areas, picture area and background image for progressive decoding.

11. A method of decoding a compressed document image stream, the method comprising:

15 extracting one or more picture areas from the compressed document image stream;

extracting one or more character areas from the compressed document image stream;

extracting a background image from the compressed data image stream;

20 identifying character blocks of the character areas with reference to dynamically generated templates in the compressed document image stream;

decoding the background image utilising a wavelet based SAQ method; and

constructing a decoded document image by adding the picture areas, the character areas and the background image.

25 12. A computer readable data storage medium having stored thereon code means for instructing a computer to execute a method of encoding a document image, the method comprising:

extracting one or more picture areas from the document image;

extracting one or more character areas from the document image;

30 obtaining a background image by subtracted the image and character areas from the document image;

classifying character blocks of the character areas with reference to dynamically generated templates; and

encoding the background image utilising a wavelet based SAQ method.

13. A computer readable data storage medium having stored thereon code means for instructing a computer to execute a method of decoding a compressed document image stream, the method comprising:

5 extracting one or more picture areas from the compressed document image stream;

extracting one or more character areas from the compressed document image stream;

extracting a background image from the compressed data image stream;

10 identifying character blocks of the character areas with reference to dynamically generated templates in the compressed document image stream;

decoding the background image utilising a wavelet based SAQ method; and

constructing a decoded document image by adding the picture areas, the character areas and the background image.

15 14. A system for encoding a document image, the system comprising:

means for extracting one or more picture areas from the document image;

means for extracting one or more character areas from the document image;

20 means for obtaining a background image by subtracted the image and character areas from the document image;

means for classifying character blocks of the character areas with reference to dynamically generated templates; and

means for encoding the background image utilising a wavelet based SAQ method.

25 15. A system for decoding a compressed document image stream, the system comprising:

means for extracting one or more picture areas from the compressed document image stream;

30 means for extracting one or more character areas from the compressed document image stream;

means for extracting a background image from the compressed data image stream;

means for identifying character blocks of the character areas with reference to dynamically generated templates in the compressed document image stream;

means for decoding the background image utilising a wavelet based SAQ method; and

5 means for constructing a decoded document image by adding the picture areas, the character areas and the background image.